2

UNIVERSITY OF MARYLAND
COMPUTER SCIENCE CENTER

COLLEGE PARK, MARYLAND
20742

TR-1098
DAAG-53-76C-0138

September 1981

TEXTURE CLASSIFICATION USING AVERAGES OF LOCAL PATTERN MATCHES

Matti Pietikäinen*
Azriel Rosenfeld
Larry S. Davis

Computer Vision Laboratory
Computer Science Center
University of Maryland
College Park, MD 20742

ABSTRACT

Laws has introduced a class of texture features based on average degrees of match of the pixel neighborhoods with a set of standard masks. These features yield better texture classification than standard features based on pairs of pixels. This paper investigates simplifications of these features, and shows that their performance is not greatly affected by their exact form, and also appears to remain the same if only local match maxima are used. It also presents an alternative definition of such features based on sums and differences of Gaussian convolutions.

The support of the Defense Advanced Research Projects Agency and the U.S. Army Night Vision Laboratory under Contract DAAG-53-76C-0138 (DARPA Order 3206) is gratefully acknowledged, as is the help of Janet Salzman in preparing this paper.

---

*Permanent address: Department of Electrical Engineering, University of Oulu, Finland

1. Introduction

Texture analysis is an important aspect of many image analysis tasks. It is used to describe and discriminate among complex regions in an image that are easier to characterize statistically rather than in detail. For a general introduction to the subject, and a review of the many approaches to the quantitative characterization of textures that have been proposed, see [1].

The suggestion that a textured region can be described in terms of the values of local properties averaged over the region dates back to at least the 1960's (see [2], p. 116; [3]; and [4], p. 419). The evidence that human texture discrimination depends on differences in second-order gray level statistics [5] can be largely accounted for in terms of differences in first-order statistics (in particular, average rates of occurrence) of local features such as lines and line ends; see, e.g., [6]. In a comparative study [7], texture classification performance based on second-order gray level statistics was found to be no better than performance based on first-order statistics (in particular, means) of gray level differences. A recent study of the role of average values of local properties in texture discrimination can be found in [8].

Recently Laws [9,10] developed and investigated a set of textural properties based on average values of local properties - specifically, matches between the pixel neighborhoods and a set of standard masks. He found that these properties performed significantly better than standard properties based on pairs of pixels. The definitions of Laws' properties are summarized in Section 2, where we also show that very similar sets of properties can be defined in terms of sums and differences of Gaussian convolutions. Gaussian convolutions have been used by others in modeling feature detection processes in human vision; see [11,12]. Section 3 presents experimental results on the performance of Laws' properties (in comparison with standard ones), and showing that their performance remains about the same if we simplify their definition or if we use only the local match maxima. (On the role of local extrema in texture analysis see [13-15].)

2. Laws' textural properties

Laws' properties, which he called "texture energy measures", are derived from three simple vectors of length 3, $L3 \equiv (1,2,1)$, $E3 \equiv (-1,0,1)$, and $S3 \equiv (-1,2,-1)$, which represent the one-dimensional operations of center-weighted local averaging, symmetric first differencing ("edge detection"), and second differencing ("spot detection"). If we convolve these vectors with themselves or each other we obtain five vectors of length 5:

$$L5 \equiv (1,4,6,4,1) = L3*L3$$

$$S5 \equiv (-1,0,2,0,-1) = E3*E3 = L3*S3$$

$$R5 \equiv (1,-4,6,-4,1) = S3*S3$$

$$E5 \equiv (-1,-2,0,2,1) = L3*E3$$

$$W5 \equiv (-1,2,0,-2,1) = E3*S3$$

where L5 is again a local average, S5 and E5 are respectively spot and edge detectors, and R5 and W5 can be regarded as "ripple" and "wave" detectors. Sets of larger vectors can be defined by repeating this convolution process [10], but we will not use them here.

If we now convolve the row vectors of length 3 or 5 with column vectors of the same length, we obtain Laws' 3×3 or 5×5 masks. The nine 3×3 masks are shown in Table 1. It can be shown that these masks span the space of 3×3 neighborhoods, i.e., any 3×3 array is a linear combination of them. For brevity, we omit the convolution and transpose signs from now on; e.g., we denote $L3^t*S3$ by L3S3.

To use these masks to describe the texture in a (sub)image, we convolve them with the image and use statistics of the results as textural properties. Laws studied the power of these convolutions, in conjunction with various statistics, to discriminate textures. Based on these studies, he concluded that the most useful 5×5 masks were the zero-sum masks obtained from L5, S5, R5, and E5 - particularly, masks such as those shown in Table 2 and/or their 90° rotations, where applicable. He also concluded that the most useful statistics were the sums of the squared or absolute values of the image after these masks are convolved with it. The sum of squares justifies the terminology "texture energy measures", but the sum of absolute values is preferable because it is computationally cheaper.

In the experiments reported in the next section, we used the 5×5 Laws features listed above, as well as the eight zero-sum 3×3 features (i.e., all but L3L3).

Masks very similar to the Laws masks can be obtained by using approximations to simple combinations of Gaussians. For example, the "five-part field" of [11] has a central excitatory zone flanked by two inhibitory zones which are in turn flanked by two weak excitatory zones, where the relative weights of the Gaussians were taken to be 0.2,-0.7,1,-0.7, and 0.2. If we scale these values by a factor of 6, we obtain (1.2,-4.2,6,-4.2, 1.2) - almost exactly the values in Laws' R5 vector. Similarly, the "three-part field" of [11] used weights of -0.5,1, and 0.5, exactly proportional to Laws' S3 vector.

3. Experiments

In most of our experiments we used four samples of each of seven textures - G=grass, R=raffia, S=sand and W=wool from Brodatz's album [16], which were also used by Laws ([10], p. 45), and three geological terrain types (L=lower Pennsylvanian shale, M=Mississippian limestone and shale, P=Pennsylvanian sandstone and shale) which were used in [7]. These samples are shown in Figure 1. To avoid effects of unequal contrast on the feature values, each sample was transformed to give it a flat gray level histogram.

Each of the masks that were tested was convolved with each of the 28 images, and the sum of absolute values was then computed. For each mask, this yields a 28-point scatter plot. Six thresholds were chosen, by hand, to separate the seven classes in this plot as completely as possible. The number of samples that could be correctly classified in this way was taken to be the score of the given mask. Table 3 shows these scores for the eight zero-sum 3×3 Laws masks, as well as for the four 5×5 masks which Laws regards as best.

For comparison, the scores of four standard texture features are also shown in Table 3. Two of these, CONX and CONY, are Haralick's [1] "contrast" feature for displacements of (1,0) and (0,1). The third feature, E/A, is "edge per unit area"; it was computed by applying the eight edge masks $\begin{matrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{matrix}$, $\begin{matrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{matrix}$,... at each pixel, taking the highest value as the gradient magnitude and the orientation of that mask as the gradient direction; suppressing nonmaxima of the gradient magnitude in the gradient direction; and counting the surviving maxima. The fourth feature, WE/A ("weighted edge per unit area"), uses the sum of the magnitudes at the maxima in place of the number of maxima. We see that these features perform consistently more poorly than the Laws features.

The best Laws features correctly classify 25 out of the 28 samples, or nearly 90%, a remarkable result for a one-feature, seven-class task. Figure 2 shows, for one sample each of the raffia and sand classes, the results of convolving the sample with each of the eight 3×3 Laws masks and scaling the results to the range [0,63].

Does the performance of the Laws features depend on the quantitative definitions of the masks, or just on their general forms? To test this, we tried six modifications of the R5R5 feature, as well as two modifications of the E5L5 feature, as shown in Table 4. Note that all of these modifications have the zero-sum property. We see from Table 3 that some of these modifications perform about as well as, and in one case even better than, the original Laws features. Thus we see that simple concentric spot masks, or symmetric, centrally weighted edge masks, may provide results comparable to those obtained from the more complex Laws masks. Laws himself found that some "ad hoc masks" did quite well; see [10] pp. 101-110.

In many cases, it may be sufficient to use the local maxima of the responses to the masks, rather than the responses at every pixel. To illustrate this, we applied local nonmaximum suppression to the outputs of three 5×5 Laws masks, R5R5, E5L5, and L5S5. Nonmaxima were suppressed in all directions for the isotropic mask R5R5, in the vertical direction for E5L5, and in the horizontal direction for L5S5 (see Table 2), to a distance of 1 or 2 pixels. For R5R5 and E5L5 suppression to distance 1 actually improved the results (to 26 and 24 correct, respectively), while suppression to distance 2 degraded them again (25 and 22 - no net change and a net reduction of 1, respectively); and for L5S5, suppression to both distances yielded a net reduction of 1 (21 correct). These results suggest that the maxima of the responses may contain the key information for texture description. Figure 3 shows the results of convolving the same two raffia and sand samples with the R5R5, E5L5, and L5S5 masks and then suppressing non-maxima out to distance 1.

A supplemental experiment was carried out with two of the classes, L and M, which have been found hard to discriminate using other types of texture features (e.g., [7]). We used either 16 64×64 images or 64 32×32 images from each class, obtained by subdividing a 256×256 image of each terrain type (Figure 4), and we used only the L5S5 feature and modifications of it. Using L5S5, the score for the 32 64×64 samples was 29 correct, i.e., over 90%, and that for the 128 32×32 samples was 101 correct (under 80%). (For the three classes L, M and P, the best result using a single texture feature in [7] was only about 70%.) Using only local maxima, the scores were 30 and 102, a slight improvement, confirming the results in the preceding paragraph. Incidentally, the relative positions of these local maxima do not seem to contain any information useful for distinguishing the textures. Figure 5 shows the positions of the maxima, and Table 5 shows cooccurrence matrices for maxima (1's) and nonmaxima (0's) for unit displacements in the four principal directions; they are nearly identical.

4. Concluding remarks

Laws' "texture energy measures", based on 3×3 or 5×5 masks, are more powerful than measures based on pairs of pixels. Their power depends on the general forms of the masks (edge-like, spot-like, etc.) rather than on the specific numerical values used in the masks, and it seems to depend primarily on the local maxima of the mask matches.

References

1. R. M. Haralick, Statistical and structural approaches to texture, Proc. IEEE 67, 1979, 786-804.

2. A. Rosenfeld, Picture Processing by Computer, Academic Press, New York, 1969.

3. J. K. Hawkins, Textural properties for pattern recognition, in B. S. Lipkin and A. Rosenfeld, eds., Picture Processing and Psychopictorics, Academic Press, New York, 1970, pp. 347-370.

4. A. Rosenfeld and A. C. Kak, Digital Picture Processing, Academic Press, New York, 1976.

5. B. Julesz, Visual pattern discrimination, IRE Trans. Information Theory 8, 1962, 84-92.

6. B. Julesz, E. N. Gilbert, L. A. Shepp, and H. L. Frisch, Inability of humans to discriminate between visual textures that agree in second-order statistics - revisited, Perception 2, 1973, 391-405.

7. J. S. Weszka, C. R. Dyer, and A. Rosenfeld, A comparative study of texture features for terrain classification, IEEE Trans. Systems, Man, Cybernetics 6, 1976, 269-285.

8. J. Beck, S. Prazdny, and A. Rosenfeld, A theory of textural segmentation, TR-1083, Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD, August 1981.

9. K. I. Laws, Texture energy measures, Proc. Image Understanding Workshop, Nov. 1979, 47-51.

10. K. I. Laws, Textured image segmentation, Rept. 940, Image Processing Institute, University of Southern California, Jan. 1980.

11. I. D. G. Macleod and A. Rosenfeld, The visibility of gratings: spatial frequency channels or bar detecting units, Vision Research 14, 1974, 909-916.

12. D. Marr, T. Poggio, and S. Ullman, Bandpass channels, zero-crossings, and early visual information processing, J. Opt. Soc. Amer. 69, 1979, 914-916.

13. S. W. Zucker, A. Rosenfeld, and L. S. Davis, Picture segmentation by texture discrimination, IEEE Trans. Computers 24, 1975, 1228-1233.

14. O. R. Mitchell, C. R. Myers, and W. Boyne, A max-min measure for image texture analysis, IEEE Trans. Computers 26, 1977, 408-414.

15. L. S. Davis, S. A. Johns, and J. K. Aggarwal, Texture analysis using generalized cooccurrence matrices, IEEE Trans. Pattern Analysis Machine Intelligence 1, 1979, 251-259.

16. P. Brodatz, Textures: A Photographic Album for Artists and Designers, Dover, New York, 1966.

$L3^t*L3$:			$L3^t*E3$:			$L3^t*S3$		
1	2	1	-1	0	1	-1	2	-1
2	[illegible]	2	-2	0	2	-2	4	-2
1	2	1	-1	0	1	-1	2	-1
$E3^t*E3$:			$E3^t*E3$:			$E3^t*E3$:		
-1	-2	-1	1	0	-1	1	-2	1
0	0	0	0	0	0	0	0	0
1	2	1	-1	0	1	-1	2	-1
$S3^t*L3$:			$S3^t*E3$:			$S3^t*S3$:		
-1	-2	-1	1	0	-1	1	-2	1
2	4	2	-2	0	2	-2	4	-2
-1	-2	-1	1	0	-1	1	-2	1

Table 1. The nine 3×3 Laws masks.

L5E5:

-1	-2	0	2	1
-4	-8	0	8	4
-6	-12	0	12	6
-4	-8	0	8	4
-1	-2	0	2	1

E5S5:

-1	0	2	0	-1
-2	0	4	0	-2
0	0	0	0	0
2	0	-4	0	2
1	0	-2	0	1

L5S5:

-1	0	2	0	-1
-4	0	8	0	-4
-6	0	12	0	-6
-4	0	8	0	-4
-1	0	2	0	-1

R5R5:

1	-4	6	-4	1
-4	16	-24	16	-4
6	-24	36	-24	6
-4	16	-24	16	-4
1	-4	6	-4	1

Table 2. Four 5×5 Laws masks judged to be most useful for texture discrimination.

3×3 Laws		5×5 Laws and classical		Modified 5×5 Laws	
Feature	Score	Feature	Score	Feature	Score
L3E3	19	E5L5	23	R5R5a	22
L3S3	23	E5S5	25	R5R5b	23
E3L3	19	L5S5	22	R5R5c	20
E3E3	21	R5R5	25	R5R5d	22
E3S3	24	CONX	20	R5R5e	24
S3L3	19	CONY	19	R5R5f	20
S3E3	24	E/A	19	E5L5a	22
S3S3	25	WE/A	19	E5L5b	24

Table 3. Number of the 28 samples correctly classified using hand-picked thresholds for eight 3×3 Laws features, four 5×5 Laws features and four classical features, and eight modified 5×5 Laws features.

R5R5a: $$\begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -4 & -4 & -4 & 1 \\ 1 & -4 & 16 & -4 & 1 \\ 1 & -4 & -4 & -4 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix}$$

R5R5b: $$\begin{matrix} 1 & 1 & 1 & 1 & 1 \\ 1 & -8 & -8 & -8 & 1 \\ 1 & -8 & 48 & -8 & 1 \\ 1 & -8 & -8 & -8 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{matrix}$$

R5R5c: $$\begin{matrix} -1 & -1 & -1 & -1 & -1 \\ -1 & -4 & -4 & -4 & -1 \\ -1 & -4 & 48 & -4 & -1 \\ -1 & -4 & -4 & -4 & -1 \\ -1 & -1 & -1 & -1 & -1 \end{matrix}$$

R5R5d: $$\begin{matrix} -2 & -2 & -2 & -2 & -2 \\ -2 & 0 & 0 & 0 & -2 \\ -2 & 0 & 32 & 0 & -2 \\ -2 & 0 & 0 & 0 & -2 \\ -2 & -2 & -2 & -2 & -2 \end{matrix}$$

R5R5e: $$\begin{matrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -10 & 0 & 0 \\ 1 & -10 & 36 & -10 & 1 \\ 0 & 0 & -10 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{matrix}$$

R5R5f: $$\begin{matrix} -1 & -1 & -2 & -1 & -1 \\ -1 & -3 & -4 & -3 & -1 \\ -2 & -4 & 48 & -4 & -2 \\ -1 & -3 & -4 & -3 & -1 \\ -1 & -1 & -2 & -1 & -1 \end{matrix}$$

E5L5a: $$\begin{matrix} -1 & -1 & -1 & -1 & -1 \\ -2 & -2 & -2 & -2 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 \end{matrix}$$

E5L5b: $$\begin{matrix} -1 & -1 & -1 & -1 & -1 \\ -8 & -8 & -8 & -8 & -8 \\ 0 & 0 & 0 & 0 & 0 \\ 8 & 8 & 8 & 8 & 8 \\ 1 & 1 & 1 & 1 & 1 \end{matrix}$$

Table 4. Six modifications of the R5R5 Laws feature, and two modifications of the E5L5 feature

L			M		
(0,1)	0	1	(0,1)	0	1
0	0.515	0.131	0	0.515	0.134
1	0.131	0.223	1	0.134	0.217
(1,0)	0	1	(1,0)	0	1
0	0.289	0.357	0	0.297	0.352
1	0.354	0.000	1	0.350	0.001
(1,1)	0	1	(1,1)	0	1
0	0.349	0.297	0	0.360	0.289
1	0.294	0.060	1	0.288	0.063
(1,-1)	0	1	(1,-1)	0	1
0	0.352	0.294	0	0.357	0.292
1	0.292	0.063	1	0.291	0.060

Table 5. (Maxima, nonmaxima) cooccurrences for the L and M textures; the displacement is shown in the upper left corner of each matrix.

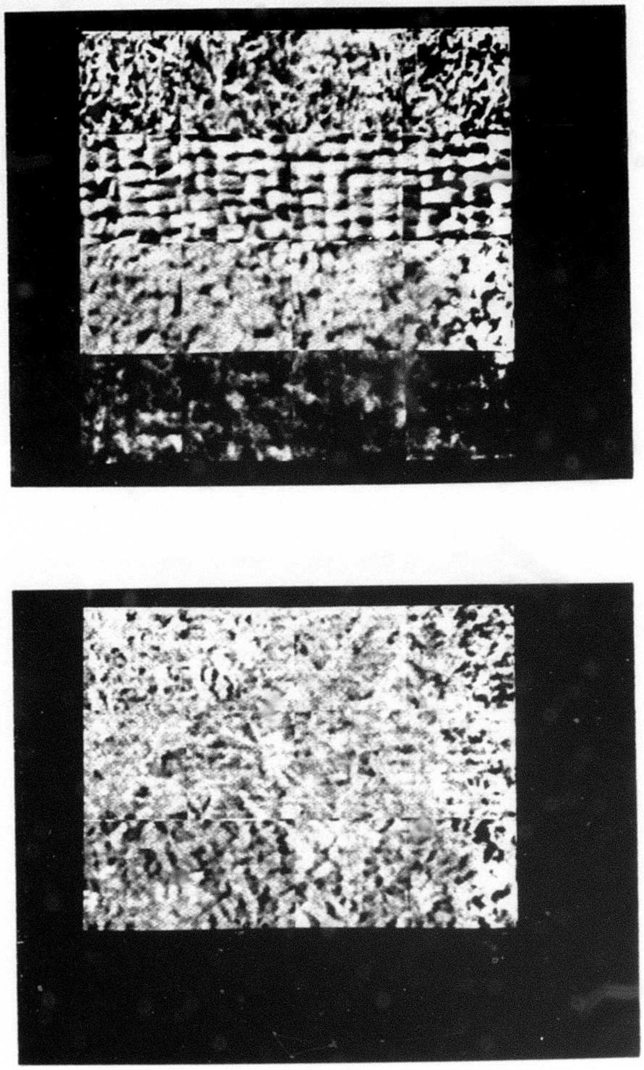


Figure 1. The 28 texture samples used in our main experiments. Top to bottom: Grass, Raffia, Sand, Wool; Lower Pennsylvanian, Mississippian, Pennsylvanian. The samples are shown prior to histogram flattening.

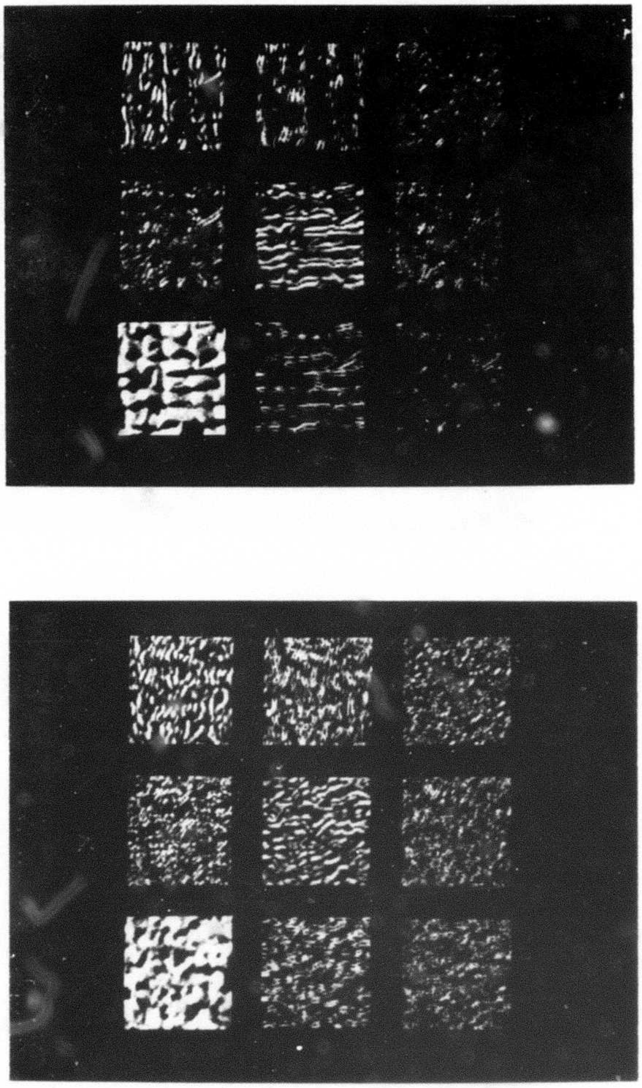


Figure 2. Results of convolving the eight 3×3 Laws masks with one each of the texture types raffia (top) and sand (bottom):

L3E3	L3S3	S3S3
E3E3	E3L3	E3S3
Original	S3L3	S3E3

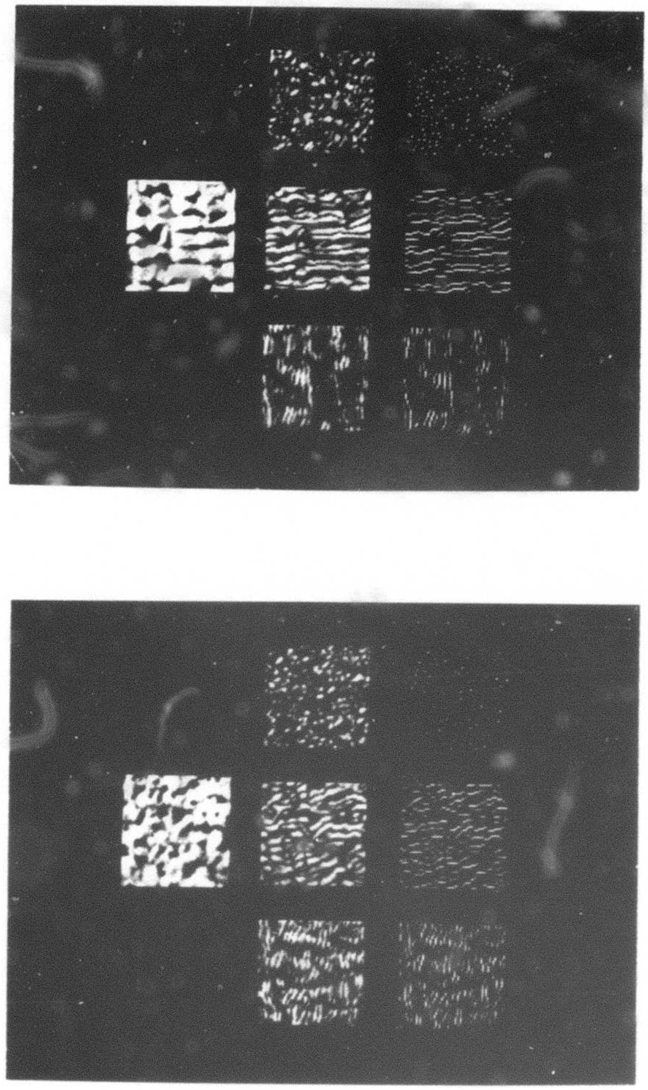
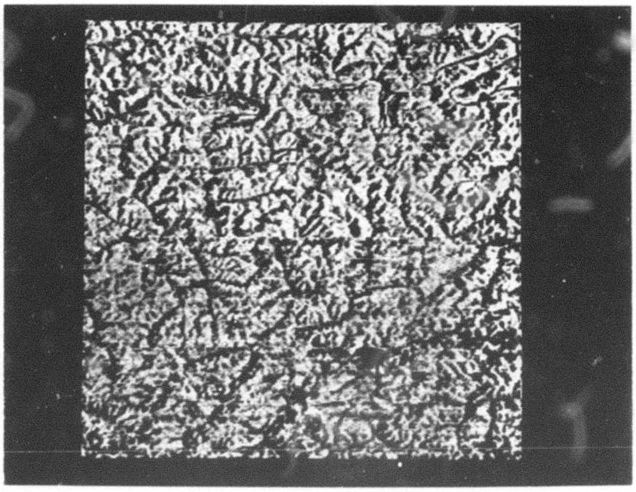


Figure 3. Results of convolving three of the 5×5 Laws masks with one sample of raffia (top) and sand (bottom), and suppressing nonmaxima out to distance 1:

	R5R5
Original	E5L5
	L5S5



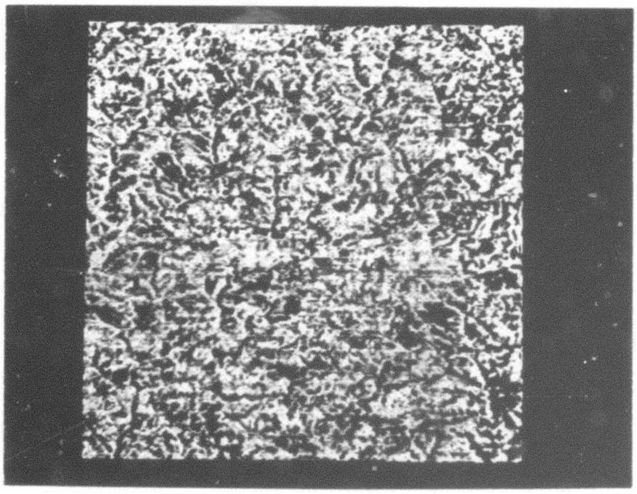
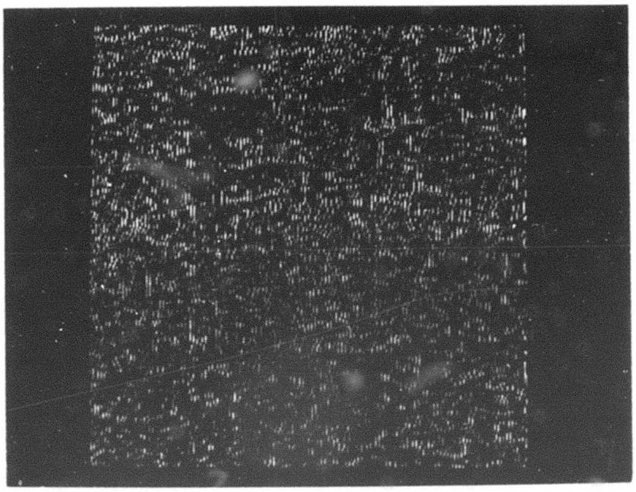


Figure 4. Lower Pennsylvanian (top) and Mississippian (bottom) images used in supplemental study. The blocks have been histogram flattened.



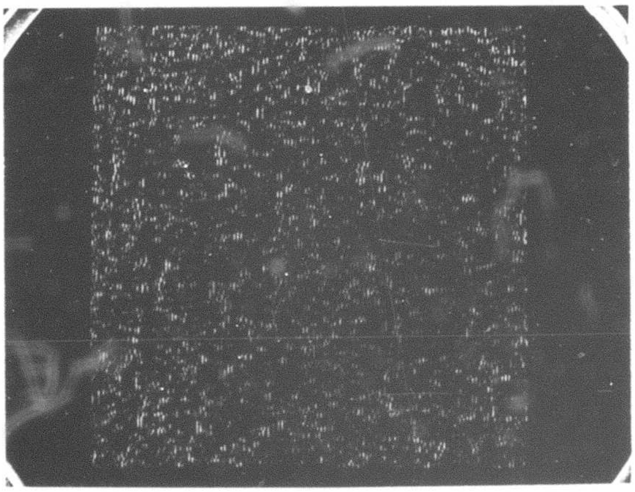


Figure 5. Positions of L5S5 maxima in the images in Figure 4.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A124 807	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TEXTURE CLASSIFICATION USING AVERAGES OF LOCAL PATTERN MATCHES		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER TR-1098
7. AUTHOR(s) Matti Pietikäinen, Larry S. Davis, Azriel Rosenfeld		8. CONTRACT OR GRANT NUMBER(s) DAAG-53-76C-0138
9. PERFORMING ORGANIZATION NAME AND ADDRESS Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD 20742		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Night Vision Laboratory, Ft. Belvoir, VA 22060		12. REPORT DATE September 1981
		13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Image processing
Pattern recognition
Texture analysis
Pattern matching

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Laws has introduced a class of texture features based on average degrees of match of the pixel neighborhoods with a set of standard masks. These features yield better texture classification than standard features based on pairs of pixels. This paper investigates simplifications of these features, and shows that their performance is not greatly affected by their exact form, and also appears to remain the same if only local match maxima are used.

It also presents an alternative definition of such features based on sums and differences of Gaussian convolutions.